

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 20:35:05 ; Search time 177 Seconds  
(without alignments)  
500.507 Million cell updates/sec

Title: US-10-736-997-134  
Perfect score: 880  
Sequence: 1 STQDGNGHGHGTHVAGTVAALN.....YASLNGTSMATPHVAGVAAL 173  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 1612378 seqs, 512079187 residues  
Total number of hits satisfying chosen parameters: 1612378  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	748	85.0	269	1	PRTM_BACSK	Q99405 bacillus sp
2	744	84.5	269	1	SUBS_BACLE	P29600 bacillus le
3	744	84.5	380	1	ELYA_BACCS	P41362 bacillus cl
4	741	84.2	380	1	ELYA_BACAO	P27693 bacillus al
5	740	84.1	269	1	SUBB_BACLE	P29599 bacillus le
6	732	83.2	382	2	Q45522	Q45522 bacillus sp
7	710	80.7	379	2	O66153	O66153 bacillus sp
8	704	80.0	378	1	ELYA_BACYA	P20724 bacillus sp
9	615	69.9	361	1	ELYA_BACHD	P41363 bacillus ha
10	615	69.9	361	2	Q45521	Q45521 bacillus sp
11	589.5	67.0	377	2	Q6L9T7	Q6L9t7 bacillus sp
12	583	66.2	374	2	Q45523	Q45523 bacillus sp
13	575	65.3	378	2	Q45466	Q45466 bacillus sp
14	562	63.9	382	1	SUBT_BACAM	P00782 bacillus am
15	560.5	63.7	379	2	Q45467	Q45467 bacillus sp
16	558	63.4	382	2	Q6IT79	Q6it79 bacillus sp
17	555	63.1	275	2	Q9R7J4	Q9r7j4 bacillus su
18	554	63.0	376	2	Q76KL9	Q76kl9 bacillus sp
19	549	62.4	382	2	O87655	O87655 bacillus su
20	545	61.9	310	2	Q9FDF2	Q9fdf2 bacillus li
21	545	61.9	374	2	Q9F941	Q9f941 bacillus li

22	545	61.9	374	2	Q9F942	Q9f942	bacillus	li
23	545	61.9	374	2	Q9F943	Q9f943	bacillus	li
24	545	61.9	379	2	Q6BCN9	Q6bcn9	bacillus	mo
25	545	61.9	379	2	Q6PNN5	Q6pnn5	bacillus	li
26	542	61.6	381	1	SUBN_BACNA	P35835	bacillus	su
27	542	61.6	381	2	Q847A2	Q847a2	bacillus	su
28	541	61.5	275	2	Q93L66	Q93l66	bacillus	su
29	541	61.5	310	2	Q9F7C2	Q9f7c2	bacillus	li
30	541	61.5	379	2	Q53521	Q53521	bacillus	li
31	539	61.3	379	2	Q45301	Q45301	bacillus	li
32	539	61.3	381	1	SUBT_BACSA	P00783	bacillus	su
33	539	61.3	381	1	SUBT_BACST	P29142	bacillus	st
34	538	61.1	379	1	SUBT_BACLI	P00780	bacillus	li
35	538	61.1	379	2	Q9FDF4	Q9fdf4	bacillus	li
36	538	61.1	379	2	Q65LP7	Q65lp7	bacillus	li
37	538	61.1	381	1	SUBT_BACSU	P04189	bacillus	su
38	537.5	61.1	404	2	Q76L29	Q76l29	bacillus	sp
39	537	61.0	310	2	Q9FDF3	Q9fdf3	bacillus	li
40	534	60.7	274	1	SUBD_BACLI	P00781	bacillus	li
41	533	60.6	275	1	SUBT_BACPU	P07518	bacillus	pu
42	530	60.2	381	2	Q84F18	Q84f18	bacillus	su
43	527	59.9	379	2	Q45300	Q45300	bacillus	li
44	511	58.1	381	2	Q7WVA6	Q7wva6	bacillus	su
45	501	56.9	379	2	Q45299	Q45299	bacillus	li
46	485.5	55.2	397	2	P97097	P97097	bacillus	sp
47	482	54.8	383	2	Q9KWR4	Q9kwr4	bacillus	pu
48	481	54.7	372	2	Q9KF13	Q9kf13	bacillus	ha
49	480	54.5	383	2	Q6SIX5	Q6six5	bacillus	pu
50	474.5	53.9	397	2	Q81DP0	Q81dp0	bacillus	ce
51	468.5	53.2	397	2	Q9S3Y3	Q9s3y3	bacillus	th
52	467.5	53.1	401	1	THES_BACSJ	Q45670	bacillus	sp
53	467.5	53.1	401	2	Q9AER8	Q9aer8	bacillus	st
54	466.5	53.0	397	2	Q63BI7	Q63bi7	bacillus	ce
55	466.5	53.0	397	2	Q81QN9	Q81qn9	bacillus	an
56	466.5	53.0	397	2	Q6HIZ4	Q6hiz4	bacillus	th
57	464.5	52.8	397	2	Q7MOW3	Q7mow3	bacillus	ce
58	459.5	52.2	397	2	Q738I4	Q738i4	bacillus	ce
59	453.5	51.5	279	1	THET_THEVU	P04072	thermoactin	
60	448.5	51.0	404	2	Q8EN89	Q8en89	oceanobacil	
61	445.5	50.6	645	1	SUBE_BACSU	P16396	bacillus	su
62	442.5	50.3	384	2	Q56365	Q56365	thermoactin	
63	435.5	49.5	565	2	Q67S71	Q67s71	symbiobacte	
64	429	48.8	196	2	Q6WFW7	Q6wfw7	bacillus	th
65	424.5	48.2	613	2	Q55477	Q55477	synechocyst	
66	423.5	48.1	488	2	Q8YY56	Q8yy56	anabaena	sp
67	423	48.1	158	2	Q6WFW6	Q6wfw6	bacillus	th
68	402	45.7	422	1	TKSU_PYRKO	P58502	pyrococcus	
69	401.5	45.6	591	2	Q65LN3	Q65ln3	bacillus	li
70	400.5	45.5	321	1	ISP_BACCS	P29140	bacillus	cl
71	396	45.0	525	2	Q9HMF2	Q9hmf2	halobacteri	
72	392.5	44.6	326	1	ISP_PAEPO	P29139	paenibacill	
73	390.5	44.4	513	2	Q64FJ8	Q64fj8	thermus	sp.
74	388.5	44.1	401	2	Q9P993	Q9p993	pyrobaculum	
75	388.5	44.1	513	1	AQL1_THEAQ	P08594	thermus	aqu
76	388.5	44.1	615	2	Q8YRA5	Q8yra5	anabaena	sp
77	388	44.1	440	2	Q9YFI3	Q9yfi3	aeropyrum	p
78	387.5	44.0	640	2	Q934J3	Q934j3	prevotella	

79	382	43.4	485	2	Q74BG6	Q74bg6 geobacter s
80	381.5	43.4	422	2	Q52407	Q52407 pseudomonas
81	378	43.0	547	2	Q8RJP4	Q8rjp4 alteromonas
82	377.5	42.9	316	2	Q6HJW5	Q6hjwt5 bacillus th
83	376	42.7	804	2	Q8RLP2	Q8rlp2 exiguobacte
84	375.5	42.7	316	2	Q739R2	Q739r2 bacillus ce
85	374	42.5	319	1	ISP1_BACSU	P11018 bacillus su
86	373.5	42.4	316	2	Q63CF9	Q63cf9 bacillus ce
87	373.5	42.4	316	2	Q81RP1	Q81rp1 bacillus an
88	373.5	42.4	323	2	Q45621	Q45621 bacillus sp
89	372.5	42.3	315	2	Q81EH9	Q81eh9 bacillus ce
90	372.5	42.3	322	2	Q9KAV3	Q9kav3 bacillus ha
91	372.5	42.3	422	2	Q9RER0	Q9rer0 vibrio mets
92	371.5	42.2	575	2	Q8PNW1	Q8pnw1 xanthomonas
93	371.5	42.2	580	1	EXPR_XANCP	P23314 xanthomonas
94	371.5	42.2	588	2	Q9KCS8	Q9kcs8 bacillus ha
95	369.5	42.0	436	2	Q45669	Q45669 bacillus sm
96	369	41.9	317	2	Q65KT7	Q65kt7 bacillus li
97	368	41.8	319	2	Q69DB4	Q69db4 bacillus sp
98	368	41.8	591	2	Q7NKC4	Q7nkc4 gloeobacter
99	367.5	41.8	358	2	Q9HF11	Q9hfl1 penicillium
100	367	41.7	1448	2	Q8YWJ8	Q8ywj8 anabaena sp

#### ALIGNMENTS

##### RESULT 1

##### PRTM\_BACSK

ID PRTM\_BACSK STANDARD; PRT; 269 AA.  
AC Q99405;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE M-protease (EC 3.4.21.-).  
OS Bacillus sp. (strain KSM-K16).  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=66692;  
RN [1]  
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
RX PubMed=15299321;  
RA Yamane T., Kani T., Hatanaka T., Suzuki A., Ashida T., Kobayashi T.,  
RA Ito S., Yamashita O.;  
RT "Structure of a new alkaline serine protease (M-protease) from  
RT Bacillus sp. KSM-K16.";  
RL Acta Crystallogr. D 51:199-206(1995).  
RN [2]  
RP SEQUENCE OF 1-23, AND CHARACTERIZATION.  
RX MEDLINE=95358832; PubMed=7632397;  
RA Kobayashi T., Hakamada Y., Adachi S., Hitomi J., Yoshimatsu T.,  
RA Koike K., Kawai S., Ito S.;  
RT "Purification and properties of an alkaline protease from alkalophilic  
RT Bacillus sp. KSM-K16.";  
RL Appl. Microbiol. Biotechnol. 43:473-481(1995).  
CC -!- COFACTOR: Binds 2 calcium ions per subunit.  
CC -!- SUBUNIT: Monomer.  
CC -!- SIMILARITY: Belongs to the peptidase S8 family.  
DR PDB; 1MPT; X-ray; @=1-269.  
DR MEROPS; S08.010; -.  
DR InterPro; IPR000209; Pept\_S8\_S53.

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 20:42:56 ; Search time 161 Seconds  
(without alignments)  
415.125 Million cell updates/sec

Title: US-10-736-997-134  
Perfect score: 880  
Sequence: 1 STQDGNGHGHGTHVAGTVAALN.....YASLNGTSMATPHVAGVAAL 173  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 1726216 seqs, 386330316 residues  
Total number of hits satisfying chosen parameters: 1726216  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : **Published Applications\_AA:\***

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep:\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	% Query	Match	Length	DB	ID	Description
1	880	100.0	173	10	US-09-824-893A-134*	Sequence 134, App	
2	880	100.0	173	16	US-10-736-997-134	Sequence 134, App	

3	872	99.1	173	10	US-09-824-893A-161*	Sequence 161, App
4	872	99.1	173	10	US-09-824-893A-164*	Sequence 164, App
5	872	99.1	173	10	US-09-824-893A-165*	Sequence 165, App
6	872	99.1	173	16	US-10-736-997-161	Sequence 161, App
7	872	99.1	173	16	US-10-736-997-164	Sequence 164, App
8	872	99.1	173	16	US-10-736-997-165	Sequence 165, App
9	868	98.6	175	10	US-09-824-893A-267	Sequence 267, App
10	868	98.6	175	16	US-10-736-997-267	Sequence 267, App
11	867	98.5	173	10	US-09-824-893A-142	Sequence 142, App
12	867	98.5	173	16	US-10-736-997-142	Sequence 142, App
13	866	98.4	173	10	US-09-824-893A-146	Sequence 146, App
14	866	98.4	173	16	US-10-736-997-146	Sequence 146, App
15	863	98.1	173	10	US-09-824-893A-145	Sequence 145, App
16	863	98.1	173	16	US-10-736-997-145	Sequence 145, App
17	862	98.0	173	10	US-09-824-893A-140	Sequence 140, App
18	862	98.0	173	16	US-10-736-997-140	Sequence 140, App
19	861	97.8	173	10	US-09-824-893A-148	Sequence 148, App
20	861	97.8	173	16	US-10-736-997-148	Sequence 148, App
21	858	97.5	173	10	US-09-824-893A-132	Sequence 132, App
22	858	97.5	173	10	US-09-824-893A-159	Sequence 159, App
23	858	97.5	173	16	US-10-736-997-132	Sequence 132, App
24	858	97.5	173	16	US-10-736-997-159	Sequence 159, App
25	857	97.4	173	10	US-09-824-893A-137	Sequence 137, App
26	857	97.4	173	16	US-10-736-997-137	Sequence 137, App
27	856	97.3	173	10	US-09-824-893A-133	Sequence 133, App
28	856	97.3	173	10	US-09-824-893A-144	Sequence 144, App
29	856	97.3	173	16	US-10-736-997-133	Sequence 133, App
30	856	97.3	173	16	US-10-736-997-144	Sequence 144, App
31	855	97.2	173	10	US-09-824-893A-138	Sequence 138, App
32	855	97.2	173	16	US-10-736-997-138	Sequence 138, App
33	852.5	96.9	172	10	US-09-824-893A-135	Sequence 135, App
34	852.5	96.9	172	16	US-10-736-997-135	Sequence 135, App
35	851	96.7	173	10	US-09-824-893A-139	Sequence 139, App
36	851	96.7	173	10	US-09-824-893A-143	Sequence 143, App
37	851	96.7	173	10	US-09-824-893A-157	Sequence 157, App
38	851	96.7	173	16	US-10-736-997-139	Sequence 139, App
39	851	96.7	173	16	US-10-736-997-143	Sequence 143, App
40	851	96.7	173	16	US-10-736-997-157	Sequence 157, App
41	850	96.6	173	10	US-09-824-893A-153	Sequence 153, App
42	850	96.6	173	10	US-09-824-893A-158	Sequence 158, App
43	850	96.6	173	16	US-10-736-997-153	Sequence 153, App
44	850	96.6	173	16	US-10-736-997-158	Sequence 158, App
45	848	96.4	173	10	US-09-824-893A-151	Sequence 151, App
46	848	96.4	173	16	US-10-736-997-151	Sequence 151, App
47	847	96.2	173	10	US-09-824-893A-150	Sequence 150, App
48	847	96.2	173	16	US-10-736-997-150	Sequence 150, App
49	847	96.2	193	10	US-09-824-893A-269	Sequence 269, App
50	847	96.2	193	16	US-10-736-997-269	Sequence 269, App
51	845	96.0	173	10	US-09-824-893A-155	Sequence 155, App
52	845	96.0	173	16	US-10-736-997-155	Sequence 155, App
53	844	95.9	173	10	US-09-824-893A-160	Sequence 160, App
54	844	95.9	173	16	US-10-736-997-160	Sequence 160, App
55	843	95.8	269	17	US-10-872-162-3 **	Sequence 3, Appli
56	843	95.8	383	17	US-10-872-162-2 **	Sequence 2, Appli
57	833	94.7	173	10	US-09-824-893A-141	Sequence 141, App
58	833	94.7	173	10	US-09-824-893A-147	Sequence 147, App
59	833	94.7	173	16	US-10-736-997-141	Sequence 141, App

60	833	94.7	173	16	US-10-736-997-147	Sequence 147, App
61	823	93.5	173	10	US-09-824-893A-154	Sequence 154, App
62	823	93.5	173	16	US-10-736-997-154	Sequence 154, App
63	814	92.5	173	10	US-09-824-893A-156	Sequence 156, App
64	814	92.5	173	16	US-10-736-997-156	Sequence 156, App
65	797	90.6	269	17	US-10-872-166-3 **	Sequence 3, Appli
66	797	90.6	383	17	US-10-872-166-2 **	Sequence 2, Appli
67	788	89.5	173	10	US-09-824-893A-168	Sequence 168, App
68	788	89.5	173	16	US-10-736-997-168	Sequence 168, App
69	781	88.8	173	10	US-09-824-893A-233	Sequence 233, App
70	781	88.8	173	16	US-10-736-997-233	Sequence 233, App
71	779	88.5	173	10	US-09-824-893A-199	Sequence 199, App
72	779	88.5	173	16	US-10-736-997-199	Sequence 199, App
73	777	88.3	173	10	US-09-824-893A-184	Sequence 184, App
74	777	88.3	173	16	US-10-736-997-184	Sequence 184, App
75	776	88.2	173	10	US-09-824-893A-236	Sequence 236, App
76	776	88.2	173	10	US-09-824-893A-247	Sequence 247, App
77	776	88.2	173	16	US-10-736-997-236	Sequence 236, App
78	776	88.2	173	16	US-10-736-997-247	Sequence 247, App
79	775	88.1	173	10	US-09-824-893A-193	Sequence 193, App
80	775	88.1	173	16	US-10-736-997-193	Sequence 193, App
81	774	88.0	173	10	US-09-824-893A-213	Sequence 213, App
82	774	88.0	173	10	US-09-824-893A-226	Sequence 226, App
83	774	88.0	173	16	US-10-736-997-213	Sequence 213, App
84	774	88.0	173	16	US-10-736-997-226	Sequence 226, App
85	774	88.0	193	10	US-09-824-893A-268	Sequence 268, App
86	774	88.0	193	16	US-10-736-997-268	Sequence 268, App
87	773	87.8	173	10	US-09-824-893A-179	Sequence 179, App
88	773	87.8	173	16	US-10-736-997-179	Sequence 179, App
89	772	87.7	173	10	US-09-824-893A-167	Sequence 167, App
90	772	87.7	173	16	US-10-736-997-167	Sequence 167, App
91	771	87.6	173	10	US-09-824-893A-166	Sequence 166, App
92	771	87.6	173	10	US-09-824-893A-232	Sequence 232, App
93	771	87.6	173	16	US-10-736-997-166	Sequence 166, App
94	771	87.6	173	16	US-10-736-997-232	Sequence 232, App
95	770	87.5	173	10	US-09-824-893A-136	Sequence 136, App
96	770	87.5	173	10	US-09-824-893A-163	Sequence 163, App
97	770	87.5	173	10	US-09-824-893A-256	Sequence 256, App
98	770	87.5	173	16	US-10-736-997-136	Sequence 136, App
99	770	87.5	173	16	US-10-736-997-163	Sequence 163, App
100	770	87.5	173	16	US-10-736-997-256	Sequence 256, App

\* Commonly-assigned US-09-824-893 issued as US Patent 6,902,922 on 18 May 2005.

\*\* 10/872,166 is not commonly assigned.

#### ALIGNMENTS

##### RESULT 1

US-09-824-893A-134

; Sequence 134, Application US/09824893A

; Publication No. US20030049827A1

; GENERAL INFORMATION:

; APPLICANT: Maxygen Inc.

; APPLICANT: No. US20030049827A1ozymes A/S

; APPLICANT: Ness, Jon

; APPLICANT: Giver, Lorraine

; APPLICANT: Cherry, Joel

; APPLICANT: Borchert, Torben

; APPLICANT: Stemmer, Willem

; APPLICANT: Minshull, Jeremy

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 20:36:50 ; Search time 39 Seconds  
(without alignments)  
426.808 Million cell updates/sec

Title: US-10-736-997-134  
Perfect score: 880  
Sequence: 1 STQDGNGHGHGTHVAGTVAALN.....YASLNGTSMATPHVAGVAAL 173  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	741	84.2	380	2	A49778	high-alkaline seri
2	732	83.2	382	2	I39780	subtilisin (EC 3.4
3	704	80.0	378	2	A33973	high-alkaline seri
4	615	69.9	361	2	G83756	subtilisin-type al
5	608	69.1	361	2	A48373	high-alkaline seri
6	583	66.2	374	2	I39781	subtilisin (EC 3.4
7	562	63.9	382	1	SUBSN	subtilisin (EC 3.4
8	542	61.6	381	2	JH0778	subtilisin (EC 3.4
9	539	61.3	381	1	SUBSS	subtilisin (EC 3.4
10	539	61.3	381	2	JQ1487	subtilisin (EC 3.4
11	538	61.1	379	1	SUBSCL	subtilisin (EC 3.4
12	538	61.1	381	1	SUBSI	subtilisin (EC 3.4
13	537	61.0	275	2	JC1085	subtilisin (EC 3.4
14	534	60.7	274	1	SUBSD	subtilisin (EC 3.4
15	533	60.6	272	2	A23624	subtilisin (EC 3.4
16	481	54.7	372	2	D83735	subtilisin-type al
17	467.5	53.1	401	2	I39974	serine proteinase
18	464.5	52.8	397	2	JW0075	cysteine-dependent
19	453.5	51.5	279	1	SUMYTV	thermitase (EC 3.4
20	445.5	50.6	645	1	SUBSMP	serine proteinase
21	442.5	50.3	384	2	JC4802	alkaline proteinas

22	424.5	48.2	613	2	S75976	hypothetical prote
23	423.5	48.1	488	2	AI1930	proteinase [import
24	400.5	45.5	321	1	S27501	alkaline proteinas
25	396	45.0	525	2	G84406	halolysin [importe
26	392.5	44.6	326	1	C41335	microbial serine p
27	388.5	44.1	401	2	A57690	aerolysin precursor
28	388.5	44.1	513	1	A35742	aqualysin (EC 3.4.
29	388.5	44.1	615	2	AH2248	proteinase [import
30	388	44.1	440	2	H72784	probable alkaline
31	374	42.5	319	2	I39866	microbial serine p
32	373.5	42.4	323	2	I39867	microbial serine p
33	372.5	42.3	322	2	G83922	intracellular alka
34	371.5	42.2	580	2	S11890	serine proteinase
35	371.5	42.2	588	2	C83836	subtilisin-type pr
36	369.5	42.0	436	2	I39973	alkaline serine pr
37	367	41.7	1448	2	AI2007	Subtilase family p
38	362.5	41.2	419	1	S25835	subtilisin (EC 3.4
39	362.5	41.2	519	2	S71451	halolysin R4 (EC 3
40	359.5	40.9	534	1	JS0173	alkaline proteinas
41	357.5	40.6	420	1	S23407	subtilisin (EC 3.4
42	353	40.1	321	2	JC5460	intracellular alka
43	347	39.4	1052	2	H83909	cell wall-associat
44	342	38.9	387	2	S11985	serine proteinase
45	340.5	38.7	536	2	JC7826	vacuolar subtilisi
46	340	38.6	384	1	SUTIKA	endopeptidase K (E
47	337	38.3	293	2	JQ0380	proteinase T (EC 3
48	337	38.3	792	2	H83736	subtilisin-type al
49	334	38.0	379	1	JC2142	alkaline proteinas
50	334	38.0	530	2	A42605	halolysin (EC 3.4.
51	334	38.0	894	2	F69730	cell wall-associat
52	331.5	37.7	533	1	JU0146	serine proteinase
53	331.5	37.7	601	2	JC4576	serine proteinase
54	327	37.2	397	2	JC7208	allergen Pen-n-13
55	326.5	37.1	467	1	S45493	serine proteinase
56	323.5	36.8	454	2	T43069	probable serine pr
57	322	36.6	388	1	S22387	cuticle-degrading
58	318.5	36.2	635	1	A29358	cerevisin (EC 3.4.
59	317.5	36.1	454	2	A75589	serine proteinase,
60	316.5	36.0	535	2	B82358	alkaline serine pr
61	313.5	35.6	478	1	S61986	subtilisin-like pr
62	312.5	35.5	715	2	JC4908	alkaline serine pr
63	307.5	34.9	591	2	A75474	serine proteinase,
64	305.5	34.7	603	2	S27055	basic serine prote
65	305	34.7	409	1	S32905	serine proteinase
66	303.5	34.5	402	1	JU0332	alkaline proteinas
67	300	34.1	442	2	A69587	intracellular alka
68	295.5	33.6	412	2	T42024	probable serine pr
69	276	31.4	1398	2	T28159	pyrolysin (EC 3.4.
70	269.5	30.6	403	1	S22184	oryzin (EC 3.4.21.
71	263	29.9	444	2	B83891	intracellular alka
72	262.5	29.8	403	1	SUASO	oryzin (EC 3.4.21.
73	262.5	29.8	403	2	JC7081	oryzin (EC 3.4.21.
74	261	29.7	393	2	T36842	probable secreted
75	257.5	29.3	491	1	S19458	hypothetical prote
76	247.5	28.1	328	2	G95392	probable protease
77	247.5	28.1	728	2	A75582	serine proteinase,
78	246	28.0	799	2	G83753	subtilisin-type pr



79	245	27.8	757	2	C84120	subtilisin-type pr
80	243	27.6	484	2	B75600	serine proteinase,
81	240.5	27.3	1433	1	A36734	bacillopeptidase F
82	238	27.0	400	2	C75336	serine proteinase,
83	237	26.9	806	2	A41341	microbial serine p
84	234	26.6	744	2	F75625	serine proteinase,
85	229	26.0	627	2	D75393	serine proteinase,
86	225	25.6	288	2	D75286	serine proteinase
87	225	25.6	4936	2	AH2515	hypothetical prote
88	216	24.5	450	2	T10926	probable serine pr
89	205.5	23.4	461	2	S23420	probable subtilisi
90	201.5	22.9	910	2	C69456	subtilisin sendai
91	197.5	22.4	454	1	A26955	alkaline serine pr
92	196	22.3	733	2	F75384	serine proteinase,
93	186	21.1	1052	2	T17093	intraluminal subti
94	185.5	21.1	995	2	C83203	probable serine pr
95	183.5	20.9	466	2	B95271	probable protease
96	183	20.8	446	2	T10024	probable signal pe
97	180	20.5	475	2	E87225	probable proteinas
98	180	20.5	1396	2	T10627	hypothetical prote
99	179.5	20.4	461	2	A70837	probable proteinas
100	178	20.2	446	2	D70597	probable signal pe

#### ALIGNMENTS

##### RESULT 1

A49778

high-alkaline serine proteinase (EC 3.4.21.-) precursor - *Bacillus alcalophilus* strain PB92)

N;Alternate names: subtilisin homolog, high-alkaline

C;Species: *Bacillus alcalophilus*

C;Date: 12-Mar-1994 #sequence\_revision 24-Feb-1995 #text\_change 09-Jul-2004

C;Accession: A49778; JC1244

R;van der Laan, J.C.; Gerritse, G.; Mulleners, L.J.S.M.; van der Hoek, R.A.C.; Quax, W.J.

Appl. Environ. Microbiol. 57, 901-909, 1991

A;Title: Cloning, characterization, and multiple chromosomal integration of a *Bacillus* alkaline protease gene.

A;Reference number: A49778; MUID:91282483; PMID:2059048

A;Accession: A49778

A;Molecule type: DNA

A;Residues: 1-380 <VAN>

A;Cross-references: UNIPROT:P27693; GB:M65086; NID:g142456; PIDN:AAA22212.1; PID:g142457

A;Experimental source: strain PB92, ATCC 31408

A;Note: amino end of mature protein confirmed by peptide sequencing

R;Takami, H.; Kobayashi, T.; Kobayashi, M.; Yamamoto, M.; Nakamura, S.; Aono, R.; Horikoshi, K.

Biosci. Biotechnol. Biochem. 56, 1455-1460, 1992

A;Title: Molecular cloning, nucleotide sequence, and expression of the structural gene for alkaline serine protease from *Alkaliphilic bacillus* sp. 221.

A;Reference number: JC1244; MUID:93043753; PMID:1368952

A;Accession: JC1244

A;Molecule type: DNA

A;Residues: 1-195,'S',197-380 <TAK>

A;Cross-references: GB:D13157; NID:g216231; PIDN:BAA02442.1; PID:g216232

A;Experimental source: *Bacillus alcalophilus* Vedder, ATCC 21522 (*Bacillus* sp. 221)

C;Superfamily: subtilisin; subtilisin homology

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 20:38:10 ; Search time 43 Seconds  
(without alignments)  
300.332 Million cell updates/sec

Title: US-10-736-997-134  
Perfect score: 880  
Sequence: 1 STQDGNGHGHGTHVAGTVAALN.....YASLNGTSMATPHVAGVAAL 173  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 513545 seqs, 74649064 residues  
Total number of hits satisfying chosen parameters: 513545  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Issued Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	841	95.6	269	4	US-09-533-480C-2	Sequence 2, Appli
2	751	85.3	269	1	US-07-706-691G-37	Sequence 37, Appl
3	751	85.3	269	1	US-08-254-021-37	Sequence 37, Appl
4	751	85.3	269	2	US-08-618-446-37	Sequence 37, Appl
5	751	85.3	269	3	US-08-980-135-37	Sequence 37, Appl
6	751	85.3	269	4	US-09-585-798-37	Sequence 37, Appl
7	749	85.1	269	1	US-07-706-691G-20	Sequence 20, Appl
8	749	85.1	269	1	US-08-254-021-20	Sequence 20, Appl
9	749	85.1	269	2	US-08-618-446-20	Sequence 20, Appl
10	749	85.1	269	3	US-08-980-135-20	Sequence 20, Appl
11	749	85.1	269	4	US-09-585-798-20	Sequence 20, Appl
12	744	84.5	269	1	US-07-706-691G-4	Sequence 4, Appli
13	744	84.5	269	1	US-07-706-691G-16	Sequence 16, Appl
14	744	84.5	269	1	US-08-254-021-4	Sequence 4, Appli
15	744	84.5	269	1	US-08-254-021-16	Sequence 16, Appl
16	744	84.5	269	1	US-08-431-387-5	Sequence 5, Appli
17	744	84.5	269	1	US-08-322-677A-10	Sequence 10, Appl
18	744	84.5	269	1	US-08-322-676-10	Sequence 10, Appl

19	744	84.5	269	2	US-08-140-083A-10	Sequence 10, Appl
20	744	84.5	269	2	US-08-618-446-4	Sequence 4, Appli
21	744	84.5	269	2	US-08-618-446-16	Sequence 16, Appl
22	744	84.5	269	3	US-08-898-218-10	Sequence 10, Appl
23	744	84.5	269	3	US-08-848-793-10	Sequence 10, Appl
24	744	84.5	269	3	US-08-980-135-4	Sequence 4, Appli
25	744	84.5	269	3	US-08-980-135-16	Sequence 16, Appl
26	744	84.5	269	3	US-09-255-502-5	Sequence 5, Appli
27	744	84.5	269	3	US-09-024-532-3	Sequence 3, Appli
28	744	84.5	269	3	US-08-090-207-2	Sequence 2, Appli
29	744	84.5	269	3	US-09-178-155-6	Sequence 6, Appli
30	744	84.5	269	3	US-09-445-270-5	Sequence 5, Appli
31	744	84.5	269	3	US-09-467-536A-1	Sequence 1, Appli
32	744	84.5	269	3	US-09-234-957-1	Sequence 1, Appli
33	744	84.5	269	4	US-09-585-798-4	Sequence 4, Appli
34	744	84.5	269	4	US-09-585-798-16	Sequence 16, Appl
35	744	84.5	269	4	US-09-417-359A-3	Sequence 3, Appli
36	744	84.5	269	4	US-09-178-173A-6	Sequence 6, Appli
37	744	84.5	269	4	US-09-512-251A-7	Sequence 7, Appli
38	744	84.5	269	4	US-09-515-150A-7	Sequence 7, Appli
39	744	84.5	269	4	US-10-075-895-1	Sequence 1, Appli
40	744	84.5	269	4	US-08-322-678-10	Sequence 10, Appl
41	744	84.5	269	4	US-09-255-501-6	Sequence 6, Appli
42	744	84.5	269	4	US-08-400-068-1	Sequence 1, Appli
43	744	84.5	269	4	US-09-196-281-10	Sequence 10, Appl
44	744	84.5	269	4	US-09-705-185-3	Sequence 3, Appli
45	744	84.5	269	4	US-09-060-854B-6	Sequence 6, Appli
46	744	84.5	269	4	US-09-177-353-6	Sequence 6, Appli
47	744	84.5	269	4	US-09-736-116-49	Sequence 49, Appl
48	744	84.5	269	4	US-10-242-549-5	Sequence 5, Appli
49	744	84.5	269	4	US-10-243-576-5	Sequence 5, Appli
50	744	84.5	269	4	US-10-033-325-6	Sequence 6, Appli
51	744	84.5	269	4	US-09-529-904-7	Sequence 7, Appli
52	744	84.5	269	4	US-09-060-872A-6	Sequence 6, Appli
53	744	84.5	269	4	US-10-310-730-1	Sequence 1, Appli
54	744	84.5	269	4	US-09-500-135C-6	Sequence 6, Appli
55	744	84.5	269	5	PCT-US95-04760-1	Sequence 1, Appli
56	744	84.5	349	1	US-07-661-378A-2	Sequence 2, Appli
57	744	84.5	380	1	US-07-918-318-2	Sequence 2, Appli
58	744	84.5	380	2	US-08-413-724-2	Sequence 2, Appli
59	744	84.5	380	3	US-08-853-494-2	Sequence 2, Appli
60	744	84.5	380	3	US-08-716-293-1	Sequence 1, Appli
61	744	84.5	380	3	US-09-445-270-7	Sequence 7, Appli
62	743	84.4	269	1	US-07-706-691G-28	Sequence 28, Appl
63	743	84.4	269	1	US-08-254-021-28	Sequence 28, Appl
64	743	84.4	269	2	US-08-618-446-28	Sequence 28, Appl
65	743	84.4	269	3	US-08-980-135-28	Sequence 28, Appl
66	743	84.4	269	4	US-09-585-798-28	Sequence 28, Appl
67	741	84.2	269	1	US-07-706-691G-5	Sequence 5, Appli
68	741	84.2	269	1	US-08-254-021-5	Sequence 5, Appli
69	741	84.2	269	1	US-08-431-387-6	Sequence 6, Appli
70	741	84.2	269	2	US-08-618-446-5	Sequence 5, Appli
71	741	84.2	269	3	US-08-980-135-5	Sequence 5, Appli
72	741	84.2	269	3	US-08-269-050-4	Sequence 4, Appli
73	741	84.2	269	3	US-08-090-207-1	Sequence 1, Appli
74	741	84.2	269	4	US-09-585-798-5	Sequence 5, Appli
75	741	84.2	269	4	US-09-512-251A-4	Sequence 4, Appli

76	741	84.2	269	4	US-09-515-150A-4	Sequence 4, Appli
77	741	84.2	269	4	US-09-196-281-7	Sequence 7, Appli
78	741	84.2	269	6	5217878-3	Patent No. 5217878
79	741	84.2	269	6	5217878-3	Patent No. 5217878
80	741	84.2	380	3	US-09-049-867-2	Sequence 2, Appli
81	741	84.2	380	3	US-08-269-050-2	Sequence 2, Appli
82	741	84.2	380	6	5217878-2	Patent No. 5217878
83	741	84.2	380	6	5336611-2	Patent No. 5336611
84	741	84.2	380	6	5217878-2	Patent No. 5217878
85	741	84.2	380	6	5336611-2	Patent No. 5336611
86	740	84.1	269	1	US-07-706-691G-18	Sequence 18, Appl
87	740	84.1	269	1	US-07-706-691G-19	Sequence 19, Appl
88	740	84.1	269	1	US-07-706-691G-21	Sequence 21, Appl
89	740	84.1	269	1	US-07-706-691G-22	Sequence 22, Appl
90	740	84.1	269	1	US-07-706-691G-23	Sequence 23, Appl
91	740	84.1	269	1	US-07-706-691G-24	Sequence 24, Appl
92	740	84.1	269	1	US-07-706-691G-25	Sequence 25, Appl
93	740	84.1	269	1	US-07-706-691G-26	Sequence 26, Appl
94	740	84.1	269	1	US-07-706-691G-29	Sequence 29, Appl
95	740	84.1	269	1	US-07-706-691G-38	Sequence 38, Appl
96	740	84.1	269	1	US-07-706-691G-40	Sequence 40, Appl
97	740	84.1	269	1	US-07-706-691G-43	Sequence 43, Appl
98	740	84.1	269	1	US-07-706-691G-46	Sequence 46, Appl
99	740	84.1	269	1	US-07-706-691G-47	Sequence 47, Appl
100	740	84.1	269	1	US-07-706-691G-49	Sequence 49, Appl

## ALIGNMENTS

## RESULT 1

US-09-533-480C-2

; Sequence 2, Application US/09533480C

; Patent No. 6777218

GENERAL INFORMATION:

; APPLICANT: Mikkelsen, Frank F.

; APPLICANT: Sejersgard, Tina Fano

10 TITLE OF INVENTION: Subtilase Enzymes

FILE REFERENCE: 6137.200-US

; CURRENT APPLICATION NUMBER: US/09/533,480C

; CURRENT FILING DATE: 2000-03-23

; NUMBER OF SEQ ID NOS: 6

```
; SOFTWARE: PatentIn version 3.1
```

; SEO ID NO 2

; LENGTH: 269

; TYPE: PRT

; ORGANISM: Artificial Sequence

FEATURE:

; OTHER INFORMATION: Synthetic

US-09-533-480C-2

Query Match 95.6%; Score 841; DB 4; Length 269;

Best Local Similarity 95.4%; Pred. No. 7.3e-73;

Matches 165; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 STQDGNGHGHVAGTVAALNNSIGVIGVAPSADLYAVKVLGANGRGSVSGIAQGLEWAAA 60

Db 55 STODGNHGHTHVAGTVAALNNSIGVIGVAPSADLYAVKVLGANGRGSVSGIAOGLEWAAA 114

Qy 61 NNMHIANMSLGSDAPSTTLGRAVNYATSQGVLVIAATGNNGSGSVGYPARYANAMAVGAT 120  
 |||  
 Db 115 NNMHIANMSLGSDAPSTTLGRAVNYATSQGVLVIAATGNNGSGSVGYPARYANAMAVGAT 174  
 Qy 121 DQNNRRANFSQYGTGIDIVAPGVNVQSTYPGNRYASLNGTSMATPHVAGVAAL 173  
 |||  
 Db 175 DQNNRRANFSQYGTGIDIVAPGVNDESTYPGSSYDSLSTSMATPHVAGVAAL 227

RESULT 2

US-07-706-691G-37

; Sequence 37, Application US/07706691G

; Patent No. 5340735

; GENERAL INFORMATION:

; APPLICANT: Goddette, Dean W

; TITLE OF INVENTION: Third Generation Protease Mutants

; NUMBER OF SEQUENCES: 104

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Henkel Corporation

; STREET: 140 Germantown Pike, Suite 150

; CITY: Plymouth Meeting

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19462

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/706,691G

; FILING DATE: 19910529

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Drach, John E.

; REGISTRATION NUMBER: 32891

; REFERENCE/DOCKET NUMBER: M4922

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-832-2215

; TELEFAX: 215-941-6067

; INFORMATION FOR SEQ ID NO: 37:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 269 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Serine Protease

; STRAIN: Bacillus lentus DSM 5843

; IMMEDIATE SOURCE:

; CLONE: S104T, S139Y, A224V

US-07-706-691G-37

Query Match 85.3%; Score 751; DB 1; Length 269;

Best Local Similarity 82.1%; Pred. No. 3.1e-64;

Matches 142; Conservative 19; Mismatches 12; Indels 0; Gaps 0;